

## Global Durum Genomic Resources Home page

In collaboration with the International Durum Wheat Genome Sequencing Consortium and the Expert Working Group on Durum Wheat Genomics and Breeding, an expert working group endorsed by Wheat Initiative, GrainGenes is making publicly available the genomic resources related to two collaborative germplasm collections: the Global Durum wheat Panel (GDP) and the Tetraploid wheat Germplasm Collection (TGC).

Both germplasm collections have been developed by Single Seed Descent, increased and genotyped with the Illumina *iSelect* wheat 90K SNP array (Wang et al. *Plant Biotechnology Journal* 12, 787-796. 2014).

**GLOBAL DURUM WHEAT PANEL (GDP):** a germplasm collection developed in the frame of the work of Expert Working Group on Durum Wheat Genomics and Breeding of the Wheat Initiative.

The GDP is a breeding dedicated tool developed with a **Forward approach**, starting from a global survey of durum wheat breeders requesting to contribute valuable genetic materials. GDP mainly focus on durum modern varieties / advanced breeding lines from worldwide, durum wheat landraces and includes a representation of wild and domesticated emmer diversity.

The GDP collection contains 1,028 accessions divided in 7 sub-groups: 1) Modern CIMMYT- and ICARDA-derived materials, semi-dwarf and vernalization-insensitive lines mostly adapted to the Mediterranean environment (288 accessions). 2) Elite semi-dwarf durum wheat lines with photoperiod and/or vernalization sensitivity mainly developed in Canada, USA, France, Italy, and Central Europe, 3) 96 non-semi-dwarf conventional height durum wheat lines of different origins. 4) Durum wheat landraces (192). 5) Inbred lines from the Evolutionary Population (EPO) contributed by INRA (c/o Pierre Roumet and Jacques David). 6) 96 Domesticated emmer, Wild Emmer and other primitive tetraploids. 7) Additional Modern European and *Fhb1* introgression lines.

The GDP (subgroups 1 to 6) is available under Term and Conditions of standard MTA at ICARDA for research, breeding and training purposes only (contact person: Filippo Bassi, Senior Durum Wheat Breeder, F.Bassi@CGIAR.org). Subgroup n.7 is available at CREA research Centre for Genomics and Bioinformatics and UNIBO (contact person Elisabetta Mazzucotelli, [elisabetta.mazzucotelli@crea.gov.it](mailto:elisabetta.mazzucotelli@crea.gov.it), Marco Maccaferri, [marco.maccaferri@unibo.it](mailto:marco.maccaferri@unibo.it)).

The manuscript describing the GDP by Mazzucotelli et al is accepted for publication in *Frontiers in Plant Science* (2020).

- Passport data file of the complete DWRC accessions (2,503 accessions) can be found here: DWRC\_list\_2503accessions.txt
- KASP® marker dataset on the DWRC accessions can be found here: DWRC\_list\_2503accessions\_Kaspar marker.txt
- Passport data file of the complete GDP accessions (1,028 accessions) can be found here: GDP\_list\_passport\_1028\_accessions.txt
- Illumina *iSelect* 90K raw polymorphism data (42,520 SNP) of the GDP accessions (1,011 accessions) can be found here GDP\_Illumina\_90K\_SNP\_1025Geno\_44536SNP.txt

- Illumina *iSelect* 90K Raw polymorphism data filtered for Mendelian inheritance, coincidence between genetic and physical map, uniqueness (non-redundant), GDP Illumina Iselect 90K raw data (MAF-unfiltered 16,633 SNPs for 1,011 accessions) can be found here GDP\_hapmap\_mapped\_filtered\_imputed\_16333SNP\_1011accessions.txt
- GDP stratification (ADMIXTURE) analysis results can be found here GDP\_Stratification\_analysis\_1011accessions.xlsx

**TETRAPLOID WHEAT GLOBAL COLLECTION (TGC**, referred to as **GTC** in Maccaferri et al. 2019, Nature Genetics): a germplasm collection comprised of 1,856 accessions developed in the frame of the durum wheat genome sequencing project, mainly focused on tetraploid genetic resources including wild and domesticated emmer, durum wheat landraces and tetraploid subspecies. The TGC(=GTC) collection has been developed with a **Reverse approach**, starting from the analysis of Gene Bank botanical/origin data (accession's passports). TGC(=GTC) has been designed to cover most of the variation present in tetraploid wheats.

A **Tetraploid Core Collection (TCC)** is comprised of 432 accessions capturing over 95% of the GDP and TGC biodiversity as shown with the Illumina *iSelect* 90K array.

The TGC-GTC and TCC is provisionally available for research, breeding and training purposes at University of Bologna and CREA-GB in small seed amount (20 seeds/accessions) under standard MTA. TGC and TCC pure seeds will be deposited to main Germplasm Banks in 2020/2021.

The manuscript describing the Global Tetraploid Wheat Collection is published by Maccaferri et al. Nature Genetics 51, 885-895 (2019) together with the Supplemental Material. Please cite the manuscript for your publications.

- Passport data file of the **TGC(=GTC) Tetraploid Germplasm Collection** final filtered collection (1,872 accessions) can be found here: TGC-GTC\_collection\_passport\_data\_file.txt
- Passport data file of the **Tetraploid Core Collection TCC** accessions (435 accessions) can be found here: TCC\_Tetraploid\_Core\_collection\_passport\_data\_file.txt. The 435 TCC accessions have been chosen using a stratification approach based on ADMIXTURE results from both GDP and TGC collections including field observations. Genotypes were selected among those clearly classified as pure representatives of defined groups and subgroups of germplasm based on molecular data and/or relevant for agronomic/resilience/ disease resistance traits
- Illumina *iSelect* 90K raw polymorphism data (34,543 SNPs) of the TGC (=GTC) accessions can be found here TGC-GTC\_collection\_SNP90K\_dataset\_34538SNP.txt
- Illumina *iSelect* 90K polymorphism data filtered for Mendelian inheritance, coincidence between genetic and physical map TGC (=GTC) Illumina Iselect 90K raw data (23,862 SNPs) can be found here TGC-GTC\_collection\_SNP90K\_dataset\_23862MendelianMappedSNP.txt
- Illumina *iSelect* 90K polymorphism data filtered for Mendelian inheritance, coincidence between genetic and physical map, non-redundant or uniqueness TGC (=GTC) Illumina Iselect 90K raw data (17,340 SNPs) can be found here TGC-GTC\_collection\_17340K\_MendelianPrunedR2099\_SNP.txt
- Overall global population structure (Admixture) for the TGC (=GTC) can be found here TGC-GTC\_collection\_population\_structure\_Supplementary file 2.xlsx

- Detailed population structure (Admixture) for the TGC (=GTC) wild emmer wheat (WEW, *T. dicoccoides*) can be found here [TGC\\_dicoccoides\\_pop\\_structure.xlsx](#)
- Detailed population structure (Admixture) for the TGC (=GTC) domesticated emmer wheat (DEW, *T. dicoccum*) can be found here [TGC\\_dicoccum\\_pop\\_structure.xlsx](#)
- Detailed population structure (Admixture) for the TGC (=GTC) durum wheat landraces (DWL, *Triticum turgidum* ssp. *durum*, *turgidum*, *turanicum*, *polonicum*, *aethiopicum*, *carthlicum*) can be found here [TGC\\_durum\\_landraces\\_pop\\_structure.xlsx](#)
- List of tetraploid-mapped QTLs projected on the Svevo genome can be found here [TetraploidQTL\\_projected.xlsx](#)